Seminar Description, SS21

Causality and Machine Learning

Gunnar König*, Timo Freiesleben, Susanne Dandl
{gunnar.koenig, susanne.dandl}@stat.uni-muenchen.de,
timo.freiesleben@campus.lmu.de
corresponding tutor*

Course Description

Classical statistical learning excels at learning associations and making accurate predictions. However, many questions like treatment effect estimation are not associative but causal in nature. Graphical models provide a general framework for describing statistical relations between random variables and performing inference. More recently, they have become a popular framework for reasoning about causal relations.
In this seminar, we will introduce the basics of inference in graphical models, discuss how they can be employed to perform causal inference and peak into recent research on links between machine learning and causality. We therefore hold two block seminars on four topics each. The four presentations of block 1 are on foundations of causal inference, the four presentations of block 2 on connections of causal inference with modern machine learning.

General Information

Kickoff and Introduction: 14th of April, 5pm CEST

Discussion of Block 1: 28th of May, 1pm
Discussion of Block 2: 25th of June, 1pm

Every topic either belongs to Block 1 or Block 2. Every presenter from block 1 is randomly assigned to one sparring partner from block 2 and vice versa. Consequently, for every topic there is an author and a reviewer.
We move the focus from the essay/paper to the presentation and the interaction within the group. As the work done as Sparring partner counts as “Koreferat”, shorter essays are admitted by the “Prüfungsordnung”.

Mattermost invitation link:
https://lmmisld-lmu-stats-slds.srv.mwn.de/signup_user_complete/?id=kmn4cubg8trd3kf5kbeqa9iy3a
Channel: seminar_causality_s21

Overview deliverables (and deadlines) by role:

- Author:
  - Abstract and Introduction (4 weeks before the Discussion Block, until Friday 23:59 pm via Mattermost channel)
  - Paper draft (2 weeks before the Discussion Block, until Friday 23:59 pm via Mattermost channel)
  - Presentation video (1 week before the Discussion Block, until Friday 23:59 pm via Mattermost channel)
  - Final essay (2 weeks after the Discussion Block, until Friday 23:59 pm via Mattermost channel)
- Sparring partner:
  - Review Extended Abstract (1 week after receival, until Wednesday 23:59 pm, via PM or email to author and supervisor of respective topic)
  - Review Paper Draft (1 week after receival, until Wednesday 23:59 pm, via PM or email to author and supervisor of respective topic): Is the structure and content of the paper understandable? The review will be in plain text and is sent to the author and supervisor.
  - Create example exercise (4 days before DB): Create short exercise that helps assessing whether the learning goal has been achieved.
  - Hosting discussion: Collecting questions, summarizing content of video, keeping time
- Participant:
  - Active participation & discussion
  - submit one question per video
  - try solving exemplary exercise

CW	15	16	17	18	19	20	21	22	23	24	25	26	27	28
all	**Kickoff**	Latest dropout				View videos, submit questions	**Discussion Session 1**			View videos, submit questions	**Discussion Session 2**			End Lecture Time
Presenters Block 1			Deadline Abstract and Introduction		Deadline Paper Draft	Deadline Video Submission		Deadline Review Abstract and Introduction	Deadline Essay Submission,	Deadline Review Paper Draft, Example Exercise	Hosting Discussion			
Presenters Block 2				Deadline Abstract and Introduction Review		Deadline Paper Draft Review, Example exercise	Hosting Discussion Deadline Abstract and Introduction		Deadline Paper Draft	Video submission			Deadline Submission Essay	

Grading & Attendance

At the kickoff event, we will provide a high-level overview of the seminar and introduce the topics we will cover. We will then assign a topic to every one of you, while trying to accommodate your preferences. The seminar will be organized in two blocks. In every block a number of presentations will be held, followed by a general critique of the presentation and a discussion of the scientific content. Your seminar grade is determined by four components:

1. **The scientific presentation (40%)**
   The goal of your presentation is to teach your topic to your fellow students. The style of the presentation is up to you, e.g., you can use slides, do a white-board presentation, etc. The duration of the presentation should be 40 minutes followed by 30 minutes discussion. The presentation will be graded along two dimensions:
   - Clarity (70%)
   - Presentation style (30%)

   Because of the current COVID-19 situation we cannot meet in person. The presentation shall therefore be recorded and uploaded by the participants, for the other students to watch. Videos will be cut at 50 minutes sharp.

2. **An Essay (30%)**
   The goal of the essay is to write a review of the topic you have been assigned to. This essay has to be submitted as PDF to your supervisor. The length of your essay depends on the number of ECTS points you receive for this seminar. If you are a Data/Computer Science student (6 ECTS points), your essay should be 8-12 pages long (ca. 10k characters). If you are a Statistics student (9 ECTS), your essay should 10-16 pages long (ca. 15k characters). In both cases the extended abstract is included in the character count. For the essay, clarity and writing style are prioritized over length. Students for the 3 ECTS version submit a short essay (4-6 pages). For 6 and 9 ECTS, an abstract should be included at the beginning of the manuscript which is also part of the character count.

3. **Abstract, introduction and first draft (10%)**
   Abstract and Introduction: The goal of abstract and introduction is to summarize motivation (problem and new approach), key arguments and if applicable results of the paper or topic. The abstracts and introduction will be sent to all course participants upon arrival. The respective reviewer will give feedback on the overall direction and moderate the discussion after the talk.
   The abstract and introduction shall be between 300 and 1200 words long. It can be used as *starting point* for abstract and introduction of the final essay.
   First draft: Submitting the draft to your reviewer will give you the chance to get feedback on your final essay. It should elaborate the rough structure and contents of the final submission and can be in form of bullet points, but should be understandable by the sparring partner.

4. **Feedback, Session Chair and Active Participation (20%)**
   The reviews should be given in time and give constructive feedback to the respective document.
   Review Abstract and Introduction: Is the motivation and content clear? The Review will be in plain text and is sent to author and supervisor.
   Review First Draft: Is the structure and content of the paper clear and understandable? The review will be in plain text and is sent to author and supervisor.
   We expect the reviewer to have understood important concepts enough to encourage a discussion of the topic at hand. As session chair he/she will quickly summarize the contents of the talk, manage time and moderate the discussion for the respective session.

We recommend (but do not require) taking the following course on scientific writing:
https://www.coursera.org/learn/sciwrite/home/welcome

Attendance is mandatory for both, block 1 and block 2. If you miss a block or your own presentation, you need to provide a medical certificate and reschedule. If you fail to do either of the two, you fail the seminar. You also fail the seminar if you drop out later than one week after the kickoff.

Main Textbooks

Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. MIT press, 2017.
Pearl, Judea. Causality. Cambridge university press, 2009.

Topics

Block 1: Foundation of Causal Inference

- Topic 1 (Pearl 1): Structural Equation Models
- Topic 2 (Pearl 2): Linking d-separation with (conditional) independence & Causal Structure Learning
- Topic 3 (Pearl 3): do-calculus: estimating the effect of interventions & counterfactuals

Block 2: Causality and Machine Learning

- Topic 4: Rubin’s Potential Outcome Framework & Causal Trees
- Topic 5: Algorithmic Recourse
- Topic 6: Causal Interpretation of Black-Box Models

Details

Good starting points Topic 1-3:
Pearl, Judea, Madelyn Glymour, and Nicholas P. Jewell. Causal inference in statistics: A primer. John Wiley & Sons, 2016.
Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. MIT press, 2017.
Pearl, Judea. Causality. Cambridge university press, 2009.
Spirtes, Peter, et al. Causation, prediction, and search. MIT press, 2000.
Hardt, Moritz and Brecht Benjamin, Patterns, predictions, and actions: A story about machine learning, 2021, https://mlstory.org/
Robert R. Tucci Introduction to Judea Pearl's Do-Calculus, 2013.https://arxiv.org/pdf/1305.5506.pdf

Topic 1: Structural Causal Models (SCMs)

Structural Causal Models are a general framework to describe data generating mechanisms. The values of variables are described as functions of exogeneous and endogeneous variables. These mechanisms can also be described with directed graphs. Typically one requires for the endogenous variables to fulfill causal sufficiency and the corresponding graph to be acyclic. Also, one requires mechanisms to be independent.
The student should convey and explain the notation on a number of examples and explain the motivation and role of the aforementioned assumptions.

Topic 2: Linking d-separation with (conditional) independence & Causal Structure Learning

Properties of the causal graph can be linked with properties of the underlying data generating distribution. More specifically, d-separation and conditional independence coincide under faithfulness and one of the Markov conditions. These are vital properties for inferring causal structure (Topic 6).
The student shall explain d-separation and conditional independence and explain and discuss the assumptions faithfulness and the global markov conditional (perfect map).
The student shall briefly sketch different approaches for causal structure learning (score-based/constraint-based algorithms) and explain one in more detail (e.g. the PC-algorithm).

Good starting points topic 2:
Malinsky, Daniel, and David Danks. "Causal discovery algorithms: A practical guide." Philosophy Compass 13.1 (2018): e12470.
Peters, Jonas, Dominik Janzing, and Bernhard Schölkopf. Elements of causal inference: foundations and learning algorithms. MIT press, 2017.

Pellet, Jean-Philippe, and André Elisseeff. "Using Markov blankets for causal structure learning." Journal of Machine Learning Research 9.Jul (2008): 1295-1342.
Constantinou, Anthony C., et al. "Large-scale empirical validation of Bayesian Network structure learning algorithms with noisy data." International Journal of Approximate Reasoning (2021).

Topic 3: Estimating the effect of interventions & counterfactuals

Given knowledge of the causal graph, the effect of interventions can be estimated under certain assumptions. The effect of interventions that can be estimated from observational data given a causal graph are called identifiable. The student shall explain under which conditions identifiability holds. Furthermore, the do-operator, which is used to distinguish interventions from normal (statistical) conditioning, shall be introduced. The front-door and back-door adjustment shall be explained.
Furthermore, as a bonus, counterfactuals in the framework of Pearl could be explained. I.e.: What is the intuition/difference to interventions? What are the three steps to compute counterfactuals in Pearl's framework.

Topic 4: Rubin's potential outcomes framework & Causal Trees

Rubin's potential outcomes framework answers how statistical models can help to draw causal inferences, i.e. how to measure the effect of causes. It is a framework to compare the outcomes we observe with the counterfactual outcomes we would have observed under a different regime or treatment.
Task 1: The student should present the framework, illustrate it on self-chosen examples and critically examine corresponding assumptions and special cases.

Most literature in the causal field focuses on estimating average treatment effects rather than identifying subgroups with heterogeneous treatment effects. But the identification of subgroups for which an intervention/treatment is particularly effective or, on the other hand, has no effect or is possibly harmful, may have important practical implications, especially in the medical context. Decision trees are a powerful method to divide observations into subgroups and there exist methods that use tree-based methods for heterogeneous treatment effect estimation.
Task 2: The student should provide an overview of tree-based methods for heterogeneous treatment effect estimation and critically examine them.

Good starting points:
Holland, Paul W. "Statistics and causal inference." Journal of the American Statistical Association 81.396 (1986): 945-960.
Jianshen Chen & Bryan Keller. "Heterogeneous Subgroup Identification in Observational Studies". Journal of Research on Educational Effectiveness. (2019). p. 578-596.
Causal Tree Learning (2020): https://www.causalflows.com/causal-tree-learning/
Susan Athey, Guido Imbens (2015): Recursive Partitioning for Heterogeneous Causal Effects. https://arxiv.org/abs/1504.01132.

Topic 5: Algorithmic Recourse

Causality is also relevant in the field of Interpretable Machine Learning. There, so-called counterfactual explanations shall guide stakeholders affected by model predictions to derive

actions that would change the algorithms outcome. This so-called algorithmic recourse is only sensibly possible while respecting the causal structure of the data.
The student shall explain the scenario and summarize and reflect the positions described in the papers below.
Starting point: *Karimi, Amir-Hossein, Bernhard Schölkopf, and Isabel Valera. "Algorithmic Recourse: from Counterfactual Explanations to Interventions." arXiv preprint arXiv:2002.06278 (2020).*
Karimi, Amir-Hossein, et al. "A survey of algorithmic recourse: definitions, formulations, solutions, and prospects." arXiv preprint arXiv:2010.04050 (2020).
Karimi, Amir-Hossein, et al. "Algorithmic recourse under imperfect causal knowledge: a probabilistic approach." arXiv preprint arXiv:2006.06831 (2020).

Topic 6: Causal Interpretation of Black-Box Models

In many scenarios, users of interpretable machine learning may be tempted to interpret the effects of interventions on the model as providing insight about the effect of an intervention on the underlying data generating mechanism. While standard IML methods are not able to provide such insight, recent proposals suggest the incorporation of knowledge about the underlying causal graph into the model interpretation, thereby enabling to reason about the effect of interventions on the data level.
The student shall motivate the need for causal variants of feature effects methods and explain the approaches in the papers below.
Starting points:
Zhao, Qingyuan, and Trevor Hastie. "Causal interpretations of black-box models." *Journal of Business & Economic Statistics* 39.1 (2021): 272-281.
Weichwald, Sebastian, et al. "Causal interpretation rules for encoding and decoding models in neuroimaging." *Neuroimage* 110 (2015): 48-59.
König, Gunnar et al, A Causal Perspective on Challenges for AI in Precision Medicine, 2019, PMBC, https://koenig.page/pdf/koenig2019_pmbc.pdf

More material:

https://cdn1.sph.harvard.edu/wp-content/uploads/sites/1268/2020/02/ci_hernanrobins_21feb20.pdf
https://arxiv.org/pdf/1305.5506v1.pdf
https://www.coursera.org/specializations/probabilistic-graphical-models
https://www.pnas.org/content/pnas/113/27/7353.full.pdf
https://causalai.net/r60.pdf

Supervision:

Timo Freiesleben: Topic 1, 5
Susanne Dandl: Topic 4
Gunnar König: Topic 2, 3, 6

Scientific Writing Rules

We emphasize that we expect you to adhere to standard rules of scientific writing. Meaning that any intellectual property that you did not create yourself has to be cited appropriately. Citing does not imply that text can be copied or copied with minor changes. You can use direct quotes if you want to preserve the wording from a reference, but that has to be tagged using parentheses “”. But after all, you are expected to write your own text, so make use of direct quotes only if absolutely necessary. Furthermore, you are not allowed to just copy figures. If you do, you have to make clear that you did and where the figure was taken from.
In the light of current events we will run every work through a plagiarism checker.

Resources on Good Scientific Writing

Coursera Course "Writing in the sciences"
Zack Lipton blog post